GenCore version 4.5

OM protein - protein search, using sw model

Run on: December 27, 2001, 16:56:13 ; Search time 16.97 Seconds
(without alignments)
505.573 Million cell updates/sec

Title: US-09-830-647-2
Perfect score: 1206
Sequence: 1 MNSGAMRIHSKGHFQGGIQV.....LKKPFVKVEDMSQSPAVHLM 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	100.5	8.3	615	1	YKZ9_YEAST	P36115 saccharomyc
2	94.5	7.8	1436	1	DPO3_STAAU	Q53665 staphylococ
3	91	7.5	908	1	YN61_YEAST	P42839 saccharomyc
4	91	7.5	994	1	PPOL_DROME	P35875 drosophila
5	90.5	7.5	584	1	BCAS_HUMAN	O75363 homo sapien
6	89	7.4	650	1	DNAK_BURPS	O68191 burkholderi
7	88	7.3	1062	1	CERU_MOUSE	Q61147 mus musculu
8	88	7.3	1214	1	BRF3_HUMAN	Q9uld4 homo sapien
9	87.5	7.3	885	1	YDGH_BACSU	P96706 bacillus su
10	86.5	7.2	1462	1	TOP2_PEA	O24308 pisum sativ
11	85	7.0	602	1	TP6B_ARCFU	O29605 archaeoglob
12	84.5	7.0	299	1	SM30_RAT	Q03336 rattus norv
13	84	7.0	815	1	RPGR_HUMAN	Q92834 homo sapien
14	84	7.0	819	1	ADVL_HUMAN	O75366 homo sapien
15	84	7.0	1059	1	CERU_RAT	P13635 rattus norv
16	83	6.9	650	1	DNAK_BURCE	P42373 burkholderi
17	83	6.9	1176	1	YOH8_YEAST	Q08236 saccharomyc
18	82.5	6.8	434	1	GATA_METJA	Q58560 methanococc
19	82.5	6.8	482	1	ODP2_YEAST	P12695 saccharomyc
20	82.5	6.8	592	1	ABP1_YEAST	P15891 saccharomyc
21	82.5	6.8	1225	1	SMC1_YEAST	P32908 saccharomyc
22	82	6.8	297	1	RRPP_RABVP	P06747 rabies viru
23	82	6.8	304	1	MMGL_MOUSE	P49300 mus musculu
24	82	6.8	320	1	YWJD_BACSU	P45864 bacillus su
25	82	6.8	322	1	RLUC_HAEIN	P44433 haemophilus
26	82	6.8	686	1	MYBB_CHICK	Q03237 gallus gall
27	81.5	6.8	299	1	SM30_MOUSE	Q64374 mus musculu
28	81.5	6.8	379	1	VLF1_NPVAC	Q06687 autographa
29	81.5	6.8	706	1	YK70_YEAST	P36166 saccharomyc
30	81.5	6.8	1131	1	AC15_MOUSE	P35601 mus musculu
31	81	6.7	622	1	PPID_HAEIN	P44092 haemophilus
32	81	6.7	700	1	GUNA_PAELA	P29719 paenibacill
33	81	6.7	1023	1	PIB4_BOVIN	Q07722 bos taurus
34	81	6.7	1478	1	BCK1_YEAST	Q01389 saccharomyc
35	81	6.7	4451	1	GRSB_BACBR	P14688 b gramicidi
36	80.5	6.7	224	1	Y945_METJA	Q58355 methanococc
37	80.5	6.7	388	1	VE2_HPV27	P36789 human papil
38	80.5	6.7	434	1	YKI2_YEAST	P36080 saccharomyc
39	80.5	6.7	673	1	TF2B_METJA	Q58192 methanococc
40	80.5	6.7	906	1	CENC_MOUSE	P49452 mus musculu
41	80.5	6.7	1523	1	SON_HUMAN	P18583 homo sapien
42	80	6.6	328	1	SWI6_SCHPO	P40381 schizosacch
43	80	6.6	726	1	BPL1_HUMAN	P50747 h biotin--p
44	80	6.6	1022	1	PIB4_HUMAN	Q15147 homo sapien
45	79.5	6.6	312	1	KIME_METJA	Q58487 methanococc

ALIGNMENTS

RESULT 1
YKZ9_YEAST
ID YKZ9_YEAST STANDARD; PRT; 615 AA.
AC P36115;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 68.9 KDA PROTEIN IN YPT52-DBP7 INTERGENIC REGION.
GN YKR019C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Duesterhoeft A., Moestl D., Poehlmann R., Philipssen P.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: TO YEAST YJL083W.
CC ---

CC ---
DR EMBL; Z28244; CAA82091.1; -.
DR PIR; S38088; S38088.
DR SGD; S0001727; YKR019C.
DR InterPro; IPR000261; EPS15_repeat.
DR SMART; SM00027; EH; 1.
KW Hypothetical protein.
SQ SEQUENCE 615 AA; 68879 MW; 242EBE45978068FD CRC64;

Query Match 8.3%; Score 100.5; DB 1; Length 615;
Best Local Similarity 20.8%; Pred. No. 1;
Matches 60; Conservative 41; Mismatches 121; Indels 67; Gaps 9;

```
Qy      3 SGAMRIHSKGHFQGGIQVKNEKNRPS-----------LKSLKTDNRPEKSKCKPLWGKV 50
          :| : |||:   : |    ||   |              | |  :  |   |  :|   |
Db    150 NGLLGIHSRSSSSRNG---SNESLTPGQRTPDNRSQENLLTSFSSGRRLSSSSMEPATNKD 206

Qy     51 FYLDLPSVTISEKLQKDIKDLGGRVE-----------EFLSKDISYLISNKKEAKFAQT 98
             ||    |  ||  :   |   |              |: | | :|||: :   :||
Db    207 SNKALPKRRPSPPLQSSLVGSGQLHENENLSSISIDSRHSLNPDTSDVISNRSQTSLSQT 266

Qy     99 LGRISPVPSPESAYTAETTSPHPSHDGSSFKS--PDTVCLSRGKLLVEK---------- 145
          : ::|   |  |  :: ||:   |: ||   :  |:     | | || |
Db    267 INQLSLCESEPSIASSNTTTT-TSNQGSGLPNLVPNYSSDMRKKKLVNKFKRKVFGSKPK 325

Qy    146 ------------AIKDHDFIPS---------NSILSNALSWGVKILHIDDIRYYIEQKK 183
                       :  |:  ||          |: :|| :     ||  ::||
Db    326 HLSSQYEMDASSEELGQHEQQPSMRFKTTLRKTSVSTNAENDHASSLHEGNLRYKY-NPS 384
```

```
Qy    184 KELYLLKKSSTSVRDGGKRVGSGAQKTRTGRLKKPFVKVEDMSQSPAVH 232
          : | :   : |  :  :   :  :  :  |:|:   |: : :   | |
Db    385 NDTYDVYDDTDSDSESDQNQDALTKPRKRDRIKR---KIRNSANKTAHH 430


RESULT   2
DPO3_STAAU
ID   DPO3_STAAU     STANDARD;      PRT;  1436 AA.
AC   Q53665; Q57110; Q9F1J9;
DT   15-DEC-1998 (Rel. 37, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   DNA POLYMERASE III POLC-TYPE (EC 2.7.7.7) (POLIII).
GN   POLC.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP8;
RX   MEDLINE=96084951; PubMed=7489915;
RA   Pacitti D.F., Barnes M.H., Li D., Brown N.C.;
RT   "Characterization and overexpression of the gene encoding
RT   Staphylococcus aureus DNA polymerase III.";
RL   Gene 165:51-56(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Inoue R., Kaito C., Tanabe M., Kamura K., Akimitsu N., Sekimizu K.;
RT   "Genetic identification of two distinct DNA polymerases, DnaE and
RT   PolC, essential for chromosomal DNA replication in Staphylococcus
RT   aureus.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED FOR REPLICATIVE DNA SYNTHESIS. THIS DNA
CC       POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC       N PYROPHOSPHATE + DNA(N).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. POLC
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D86727; BAA13160.1; -.
DR   EMBL; Z48003; CAA88043.1; -.
DR   EMBL; AB053353; BAB20885.1; -.
DR   InterPro; IPR000520; Exonuclse.
DR   InterPro; IPR003141; PHP_N.
DR   Pfam; PF00929; Exonuclease; 1.
DR   Pfam; PF02231; PHP_N; 1.
DR   SMART; SM00479; EXOIII; 1.
DR   SMART; SM00481; POLIIIAc; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
KW   Nuclease; Exonuclease.
FT   DOMAIN      420    584       EXONUCLEASE.
FT   CONFLICT    207    208       NE -> KQ (IN REF. 1).
FT   CONFLICT    952    952       T -> K (IN REF. 1).
FT   CONFLICT   1030   1049       MISSING (IN REF. 1; CAA88043).
FT   CONFLICT   1035   1035       R -> A (IN REF. 1).
FT   CONFLICT   1147   1153       EFGTGFV -> NSDRIR (IN REF. 1).
FT   CONFLICT   1260   1260       M -> I (IN REF. 1).
FT   CONFLICT   1409   1409       S -> Y (IN REF. 1).
SQ   SEQUENCE   1436 AA;  162459 MW;  2F70E034C0FBF723 CRC64;

  Query Match             7.8%;  Score 94.5;  DB 1;  Length 1436;
  Best Local Similarity   25.1%;  Pred. No. 9.4;
  Matches   57;  Conservative   32;  Mismatches   87;  Indels   51;  Gaps   12;

Qy     18 IQVKNEKNRPSLKSLKTDNRP-EKSKCKPLWGKVFYLDLPSVTISEKLQKDIKDLGGRVE 76
          : :|:|   | ::    : |    :  : |:|:    ||| : | ::|:|::| : |
Db    805 VPIKDELYTPRMEGANEEIRELSYANARKLYGE----DLPQIVI-DRLEKELKSIIGN-- 857

Qy     77 EFLSKDISYLISN---KKEAKFAQTLGRISPVPSPESAYTAETT--SPHPSH-------- 123
             : ||||    ||       :|    | |    |   | |  :| | |
Db    858 ---GFAVIYLISQRLVKKSLDDGYLVGSRGSVGSSFVATMTEITEVNPLPPHYICPNCKT 914

Qy    124 --------DGSSFKSPDTVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDI 175
                   || |  ||  | : |  |    ||:   ||  : |      | |:  | |:
Db    915 SEFFNDGSVGSGFDLPDKTCETCGAPL----IKEGQDIPFETFLGFK---GDKVPDI-DL 966

Qy    176 RYYIEQKKKELYLLKKSSTSVRDGGKRV------GSGAQKTRTGRLK 216
          :  | :         : | |  |  :|       |: |:||  | :|
Db    967 NFSGEYQPN-----AHNYTKVLFGEDKVFRAGTIGTVAEKTAFGYVK 1008


RESULT   3
YN61_YEAST
ID   YN61_YEAST     STANDARD;      PRT;   908 AA.
AC   P42839;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL 102.5 KDA PROTEIN IN KRE1-HXT14 INTERGENIC REGION.
GN   YNL321W OR N0339.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1676;
RX   MEDLINE=96076632; PubMed=7502583;
RA   Maftahi M., Nicaud J.-M., Levesque H., Gaillardin C.;
RT   "Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV
RT   identifies six known genes, a new member of the hexose transporter
RT   family and ten new open reading frames.";
RL   Yeast 11:1077-1085(1995).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: TO S.POMBE SPAC521.04C.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z46259; CAA86376.1; -.
DR   EMBL; Z71597; CAA96252.1; -.
DR   SGD; S0005265; YNL321W.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM    245    265       POTENTIAL.
FT   TRANSMEM    409    429       POTENTIAL.
FT   TRANSMEM    495    515       POTENTIAL.
FT   TRANSMEM    531    551       POTENTIAL.
FT   TRANSMEM    561    581       POTENTIAL.
FT   TRANSMEM    588    608       POTENTIAL.
FT   TRANSMEM    627    647       POTENTIAL.
FT   TRANSMEM    687    707       POTENTIAL.
FT   TRANSMEM    747    767       POTENTIAL.
FT   TRANSMEM    784    804       POTENTIAL.
FT   TRANSMEM    817    837       POTENTIAL.
FT   TRANSMEM    852    872       POTENTIAL.
FT   TRANSMEM    886    906       POTENTIAL.
SQ   SEQUENCE   908 AA;  102498 MW;  180E0EEC77DC111E CRC64;
```

```
  Query Match             7.5%;  Score 91;  DB 1;  Length 908;
  Best Local Similarity   31.1%;  Pred. No. 9.8;
  Matches   41;  Conservative   19;  Mismatches   42;  Indels   30;  Gaps    9;

Qy      32 LKTDNRPEKSKCKP------LW--GKVFYLDLPSVTISEKLQKDIKDLGGRVEEFLSKDI 83
           | |::  | |||         ||  ||: ||   :   : ||:| || |  :::| :
Db     271 LVTNSAKEYSKCLYKLANYFLWPFGKMVYL----LQDEQYLQED-KDEGISMQQFYNWVT 325

Qy      84 SY---LISNKKEAKFAQTLGRISPVPSPESAY---TAETT-----SPHPSHDGSSFKSPD 132
           ||   |: :: :||| |      | |: ||:     | ||     | |||::    : |
Db     326 SYSNRLVFHQSQAKFQQR--EDHPAPATESSSLMPPANTTATPLNSNHPSYNSIRHEIPH 383

Qy     133 TVC----LSRGK 140
                  |||
Db     384 AAAQRRYFGRGK 395


RESULT   4
PPOL_DROME
ID   PPOL_DROME     STANDARD;      PRT;   994 AA.
AC   P35875; Q9W5Q5; Q9W5S1;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   POLY [ADP-RIBOSE] POLYMERASE (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   RIBOSYLTRANSFERASE) (POLY[ADP-RIBOSE] SYNTHETASE).
GN   PARP OR CG17696/CG17718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93234521; PubMed=8475096;
RA   Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA   Sugimura T., Miwa M.;
RT   "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RT   leucine zipper in the auto-modification domain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN   [2]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RC   STRAIN=CANTON-S;
RX   MEDLINE=98234380; PubMed=9565614;
RA   Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT   "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RT   distribution of its mRNA during development.";
RL   J. Biol. Chem. 273:11881-11886(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-RIBOSYL)[N]-ACCEPTOR =
CC       NICOTINAMIDE + (ADP-D-RIBOSYL)[N+1]-ACCEPTOR.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC       OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC       IN LATER EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC       EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D13806; BAA02964.1; -.
DR   EMBL; AF051548; AAC24518.1; -.
DR   EMBL; AF051544; AAC24518.1; JOINED.
DR   EMBL; AF051545; AAC24518.1; JOINED.
DR   EMBL; AF051546; AAC24518.1; JOINED.
DR   EMBL; AF051547; AAC24518.1; JOINED.
DR   EMBL; AE002935; AAF45400.1; -.
DR   EMBL; AE002666; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE002892; AAF45445.2; ALT_SEQ.
DR   PIR; A47474; A47474.
DR   HSSP; P26446; 4PAX.
DR   FlyBase; FBgn0010247; Parp.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
```

```
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.
FT   DNA_BIND      1    367
FT   DOMAIN      368    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      380    454       BRCT.
FT   DOMAIN      508    994       NAD-BINDING.
FT   ZN_FING      19     54       PARP-TYPE.
FT   ZN_FING     123    161       PARP-TYPE.
FT   DOMAIN      208    210       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   ACT_SITE    875    875       BY SIMILARITY.
FT   VARSPLIC    376    564       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   994 AA;  113791 MW;  ACA85A270DD29E08 CRC64;


  Query Match             7.5%;  Score 91;  DB 1;  Length 994;
  Best Local Similarity   21.2%;  Pred. No. 11;
  Matches   62;  Conservative   46;  Mismatches   80;  Indels  104;  Gaps   14;

Qy      21 KNEKNRPSLKSLKTDN---RPEKSKCKPLWGKVFYLDLPSVTIS---EKLQKDIKDLGGR 74
           |:  :|  ||: | ::   ||   :  |    :: |    : :    ::|:| |::|||:
Db     356 KSTFSRSLLKTNKNNDVLVRPTIPRISP---PLYNLKFSIIGLKNQHKELRKRIENLGGK 412

Qy      75 VEEFLSKDISYLISNKKE-------AKFAQTLG-RISPV---------------------- 105
            |  :|::   :|| : |        |||: ||  | |:
Db     413 FEVKISENTIAIISTELEIQKKSTRMKFAEELGIHIVPIEFLDFVEADTEGAIKYINSTC 472

Qy     106 -----PSPESAYTAETTSPHPSHD--------GSSFKSPDTVCLSRGKLLVEKAIKDHDF 152
                |:|    |||    |:           :||  | : :     | : |   | :
Db     473 ICSWGTDPKSRIPKETTKSLNSNSIYTKSMPVSRTFKVKDGLAVDPDSGLEDIA---HVY 529

Qy     153 IPSNSILSNALSWGVKILHIDDIR-----YY----IEQKKKELYLLKKSSTSVRDGGKRV 203
           : ||:  |        :| : ||:     ||    ::  ||| | : :|   :   |  :
Db     530 VDSNNKYS-------VVLGLTDIQRNKNSYYKVQLLKADKKEKYWIFRSWGRI---GTNI 579

Qy     204 GSG------------------AQKT------------RTGRLKKPFVKVED 224
           |:                   | ||            ||||:    :: :|
Db     580 GNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVKRTGRMYPIEIQYDD 631


RESULT   5
BCAS_HUMAN
ID   BCAS_HUMAN     STANDARD;      PRT;   584 AA.
AC   O75363;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   BREAST CARCINOMA AMPLIFIED SEQUENCE 1 (NOVEL AMPLIFIED IN BREAST
DE   CANCER 1) (AMPLIFIED AND OVEREXPRESSED IN BREAST CANCER).
GN   BCAS1 OR NABC1 OR AIBC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98337979; PubMed=9671742;
RA   Collins C., Rommens J.M., Kowbel D., Godfrey T., Tanner M.,
RA   Hwang S.-I., Polikoff D., Nonet G., Cochran J., Myambo K., Jay K.E.,
RA   Froula J., Cloutier T., Kuo W.-L., Yaswen P., Dairkee S.,
RA   Giovanola J., Hutchinson G.B., Isola J., Kallioniemi O.-P.,
RA   Palazzolo M., Martin C., Ericsson C., Pinkel D., Albertson D.,
RA   Li W.-B., Gray J.W.;
RT   "Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2
RT   and overexpressed in breast carcinoma.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:8703-8708(1998).
CC   -!- TISSUE SPECIFICITY: OVEREXPRESSED IN MOST BREAST CANCER CELL
CC       LINES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041260; AAC39896.1; -.
DR   MIM; 602968; -.
SQ   SEQUENCE   584 AA;  61719 MW;  2290CC0B37129944 CRC64;


  Query Match             7.5%;  Score 90.5;  DB 1;  Length 584;
  Best Local Similarity   26.3%;  Pred. No. 6;
  Matches   36;  Conservative   22;  Mismatches   50;  Indels   29;  Gaps    6;

Qy      23 EKNRPSLKSLKTDNRPEKSKCKPLWGKVFYLDLPSVTISEKLQKDIKDLGGRVEEFLSKD 82
           | |   :   ||   | |:: |                     || :| |.       | |
Db     282 ENNNSIMSFFKTLVSPNKAETK--------------------KDPEDTGAEKSPTTSAD 320

Qy      83 ISYLISNKKE--AKFAQTLGRISPVPSPESAYTAETTSPHPSHDGSSFKSPDTVCLSRGK 140
           :    |:|    ::  |  |: |   :| | :|   |:| |: :|:  || |  |  ||
Db     321 LK---SDKANFTSQETQGAGKNSKGCNP-SGHTQSVTTPEPAKEGTKEKSGPT-SLPLGK 375

Qy     141 LLVEKAIKDHDFIPSNS 157
           |  :|::|: | :|: :
Db     376 LFWKKSVKE-DSVPTGA 391


RESULT   6
DNAK_BURPS
ID   DNAK_BURPS     STANDARD;      PRT;   650 AA.
AC   O68191;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CHAPERONE PROTEIN DNAK (HEAT SHOCK PROTEIN 70) (HEAT SHOCK 70 KDA
DE   PROTEIN) (HSP70).
GN   DNAK.
OS   Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
OX   NCBI_TaxID=28450;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 23343;
RA   See L.H., Yap E.H., Yap E.P.H.;
RT   "Isolation and sequencing of the heat shock protein 70 (hsp70/dnaK)
RT   gene in Burkholderia pseudomallei.";
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC   -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF016711; AAC15473.1; -.
DR   HSSP; P04475; 2BPR.
DR   InterPro; IPR001023; HSP70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   PROSITE; PS00297; HSP70_1; 1.
DR   PROSITE; PS00329; HSP70_2; 1.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   Chaperone; ATP-binding; Heat shock.
SQ   SEQUENCE   650 AA;  69735 MW;  BE46330B6DE174D0 CRC64;


  Query Match             7.4%;  Score 89;  DB 1;  Length 650;
  Best Local Similarity   24.0%;  Pred. No. 9.2;
  Matches   50;  Conservative   36;  Mismatches   74;  Indels   48;  Gaps   12;
```

```
Qy      17 GIQVKNEKNRPSLKSL-----KTDNRPEKSKCKPLWGKVFYLDLPSVT-ISEKLQKDIKD 70
           |::||   | |:  :|     | :    |     | |  | : :  :  :  ::| ::
Db     162 GLEVKRIINEPTAAALAFGLDKAEKGDRKIAVYDLGGGTFDVSIIEIADVDGEMQFEVLS 221

Qy      71 ------LGGRVEEFLSKDISYLISN-KKE-----AKFAQTLGRISPVPSP-----ESAYT 113
                |||  |:|  : | |:|   |||     :|    | |:            |:
Db     222 TNGDTFLGG--EDFDQRIIDYIIGEFKKEQGVDLSKDVLALQRLKEAAEKAKIELSSSQQ 279

Qy     114 AETTSPHPSHDGSSFKSPDTVCLSRGKL------LVEKAIKDHDFIPSNSILSNALSWGV 167
           |   |: : | |  |  : : ::| ||      |||: |:     |  : : :|   ||
Db     280 TEINLPYITADASGPKHLN-LKVTRAKLEALVEDLVERTIE-----PCRTAIKDA---GV 330

Qy     168 KILHIDDIRYY--------IEQKKKELY 187
           |:  |||:           :::| || :
Db     331 KVSDIDDVILVGGQTRMPKVQEKVKEFF 358
```

```
RESULT   7
CERU_MOUSE
ID   CERU_MOUSE     STANDARD;      PRT;  1062 AA.
AC   Q61147;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE).
GN   CP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Klomp L.W.J., Farhangrazi Z.S., Choi D.W., Gitlin J.D.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=96294736; PubMed=8690795;
RA   Klomp L.W.J., Farhangrazi Z.S., Dugan L.L., Gitlin J.D.;
RT   "Ceruloplasmin gene expression in the murine central nervous system.";
RL   J. Clin. Invest. 98:207-215(1996).
CC   -!- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER
CC       MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS
CC       ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT
CC       AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: 4 FE(2+) + 4 H(+) + O(2) = 4 FE(3+) + 2 H(2)O.
CC   -!- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO
CC       THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS
CC       KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED
CC       BINUCLEAR.
CC   -!- TISSUE SPECIFICITY: MANY TISSUES, INCLUDING LIVER, EYE AND BRAIN.
CC   -!- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC       2 PLASTOCYANIN-LIKE REPEATS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49430; AAB07996.1; -.
DR   HSSP; P00450; 1KCW.
DR   MGD; MGI:88476; Cp.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR002355; MultiCu_oxidse2.
DR   Pfam; PF00394; Cu-oxidase; 3.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR   PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW   Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;
KW   Signal.
FT   SIGNAL        1     19       BY SIMILARITY.
FT   CHAIN        20   1062       CERULOPLASMIN.
FT   DOMAIN       20    356       F5/8 TYPE A 1.
FT   DOMAIN       20    199       PLASTOCYANIN-LIKE 1.
FT   DOMAIN      208    356       PLASTOCYANIN-LIKE 2.
FT   DOMAIN      369    713       F5/8 TYPE A 2.
FT   DOMAIN      369    555       PLASTOCYANIN-LIKE 3.
FT   DOMAIN      565    713       PLASTOCYANIN-LIKE 4.
FT   DOMAIN      725   1057       F5/8 TYPE A 3.
FT   DOMAIN      725    896       PLASTOCYANIN-LIKE 5.
FT   DOMAIN      904   1057       PLASTOCYANIN-LIKE 6.
FT   DISULFID    173    199       BY SIMILARITY.
FT   DISULFID    275    356       BY SIMILARITY.
FT   DISULFID    529    555       BY SIMILARITY.
FT   DISULFID    632    713       BY SIMILARITY.
FT   DISULFID    870    896       BY SIMILARITY.
FT   METAL       120    120       COPPER (TYPE 2) (BY SIMILARITY).
FT   METAL       122    122       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL       179    179       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL       181    181       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL       990    990       COPPER (TYPE 1) (BY SIMILARITY).
FT   METAL       993    993       COPPER (TYPE 2) (BY SIMILARITY).
FT   METAL       995    995       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL      1035   1035       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL      1036   1036       COPPER (TYPE 1) (BY SIMILARITY).
FT   METAL      1037   1037       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL      1041   1041       COPPER (TYPE 1) (BY SIMILARITY).
FT   METAL      1046   1046       COPPER (TYPE 1) (BY SIMILARITY).
FT   CARBOHYD    138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    583    583       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    625    625       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    757    757       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    922    922       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1062 AA;  121159 MW;  F3F52ED09A238F16 CRC64;
```

```
  Query Match             7.3%;  Score 88;  DB 1;  Length 1062;
  Best Local Similarity   21.2%;  Pred. No. 21;
  Matches   45;  Conservative   30;  Mismatches   81;  Indels   56;  Gaps   10;

Qy      10 SKGHFQGGIQVKNEKNRPSLKSLKTDNRPEKSKCKPLW----GKVFYLDLPSVTISEKLQ 65
           |  : | |     | : :|    ||    |:  || |    | |       : :|:
Db      52 SNFYLQNGPDRIGRKYKKALYFEYTDGTFSKTIDKPAWLGFLGPVI-----KAEVEDKVY 106

Qy      66 KDIKDLGGRVEEFLSKDISYLISNKKEAKFAQTLGRISP------------VPSPESAY 112
             :|:|  |:  | :  ::|     || :     | : |              :|  :  |
Db     107 VHLKNLASRIYTFHAHGVTY----TKEYE-----GAVYPDNTTDFQRADDKVLPGQQYVY 157

Qy     113 TAETTSPHP------------SH------DGSSFKSPDTVCLSRGKLLVEKAIK-DHDF 152
           | |              ||        |    |  :|  :| |  ||    | :|
Db     158 VLHANEPSPGEGDSNCVTRIYHSHVDAPKDIASGLIGPLILC-KKGSLYKEKEKNIDQEF 216

Qy     153 IPSNSILSNALSWGVKILHIDDIRYYIEQKKK 184
           :   |::   ||| ::    |:|: :  : :|
Db     217 VLMFSVVDENLSWYLE----DNIKTFCSEPEK 244
```

```
RESULT   8
BRF3_HUMAN
ID   BRF3_HUMAN     STANDARD;      PRT;  1214 AA.
AC   Q9ULD4;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   BROMODOMAIN AND PHD FINGER-CONTAINING PROTEIN 3 (FRAGMENT).
GN   BRPF3 OR KIAA1286.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
```

```
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K.-I., Kikuno R., Hirosawa M., Nomura N.,
RA   Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
CC   -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 PWWP DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB033112; BAA86600.1; -.
DR   InterPro; IPR001487; Bromodomain.
DR   InterPro; IPR001965; PHD.
DR   InterPro; IPR000313; PWWP.
DR   Pfam; PF00439; bromodomain; 1.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00855; PWWP; 1.
DR   PRINTS; PR00503; BROMODOMAIN.
DR   SMART; SM00297; BROMO; 1.
DR   SMART; SM00249; PHD; 2.
DR   SMART; SM00293; PWWP; 1.
DR   PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR   PROSITE; PS50014; BROMODOMAIN_2; 1.
DR   PROSITE; PS50812; PWWP; 1.
KW   Zinc-finger; Bromodomain.
FT   NON_TER       1      1
FT   ZN_FING     223    271       PHD-TYPE.
FT   DOMAIN      415    441       GLU-RICH.
FT   DOMAIN      615    685       BROMODOMAIN.
FT   DOMAIN     1085   1168       PWWP.
SQ   SEQUENCE   1214 AA;  136598 MW;  CA490810622109CD CRC64;
```

```
  Query Match             7.3%;  Score 88;  DB 1;  Length 1214;
  Best Local Similarity   21.5%;  Pred. No. 25;
  Matches   51;  Conservative   40;  Mismatches   80;  Indels   66;  Gaps   11;

Qy      14 FQGGIQVKNEKNRPSLKSLKTDNRPEKSKCKPLWGKVFYLDLPSVTISEKLQKDIK-DLG 72
           | | ::    |:| : :  :: : |  || |    : |          : |: :: |:
Db       6 FPGAMRKPRRKSRQNAEGRRSPS-PYSLKCSPTRETLTY---------AQAQRIVEVDID 55

Qy      73 GRV--------------EEFLSKDISYLISNKKEAKFAQTLGRISPVPSPESAYTAETTS 118
           ||:              :|  ::||:   |||: ::  |  |: |  || :     |: |
Db      56 GRLHRISIYDPLKIITEDELTAQDITECNSNKENSEQPQFPGK-SKKPSSKGK-KKESCS 113

Qy     119 PHPSHDGSSFKSP-------DTVCLSRGKLL-------VEKAIKDHDFIPSNSILSNALS 164
           | |  |:||  |       |:       |          :|| :| |    :      |:
Db     114 KHAS--GTSFHLPQPSFRMVDSGIQPEAPPLPAAYYRYIEKPPEDLDAEVEYDMDEEDLA 171

Qy     165 W-----------GVKILHIDDIRYYIEQKKKELYLLKKSSTSVRDGGKRVGSGAQKT 210
           |           |  ::  |     ::: :|| ||  :|            ||||::
Db     172 WLDMVNEKRRVDGHSLVSADTFELLVDRLEKESYLESRS-----------SGAQQS 216
```

```
RESULT   9
YDGH_BACSU
ID   YDGH_BACSU     STANDARD;      PRT;   885 AA.
AC   P96706;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PUTATIVE MEMBRANE PROTEIN YDGH.
GN   YDGH.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kasahara Y., Nakai S., Lee S., Sadaie Y., Ogasawara N.;
RT   "A 148 kbp sequence of the region between 35 and 47 degree of the
RT   Bacillus subtilis genome.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE MMPL FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB001488; BAA19398.1; -.
DR   EMBL; Z99106; CAB12372.1; -.
DR   EMBL; Z99107; CAB12384.1; -.
DR   SubtiList; BG12175; ydgH.
DR   InterPro; IPR000731; HMGCR_patched_5TM.
DR   PROSITE; PS50156; SSD; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM      9     29       POTENTIAL.
FT   TRANSMEM    181    201       POTENTIAL.
FT   TRANSMEM    202    222       POTENTIAL.
FT   TRANSMEM    227    247       POTENTIAL.
FT   TRANSMEM    278    298       POTENTIAL.
FT   TRANSMEM    304    324       POTENTIAL.
FT   TRANSMEM    354    374       POTENTIAL.
FT   TRANSMEM    716    736       POTENTIAL.
FT   TRANSMEM    740    760       POTENTIAL.
FT   TRANSMEM    772    792       POTENTIAL.
FT   TRANSMEM    817    837       POTENTIAL.
FT   TRANSMEM    847    867       POTENTIAL.
SQ   SEQUENCE   885 AA;  95488 MW;  BDCBE16179911D75B CRC64;
```

```
  Query Match             7.3%;  Score 87.5;  DB 1;  Length 885;
  Best Local Similarity   23.3%;  Pred. No. 18;
  Matches   47;  Conservative   30;  Mismatches   74;  Indels   51;  Gaps    8;

Qy       9 HSKGHFQGGIQVKNEKNRPSLKSLKTDN-------------RPEKSKCKPLWGKVFYLD 54
           : ||  |      :|:    ||    ||                :  ::: : :  |:  :|
Db      34 NQKGQAQLPADAVSERANAILKQAGEDNNSISVVFTLDNAIKKETENQLRIIIDKIKKID 93

Qy      55 -LPSVTISEKLQKDIKDLGGRVEEFLSKD-------ISYLISNKKEAKFAQTLGRISPVP 106
            :  ||     :|::||          : :|||       ::   |:||  | |  : :|  ||
Db      94 GVEEVTSPLSAEKEVKD------QLMSKDKKTVLMPVTITGSDKKAEKIADEIYQI--VP 145

Qy     107 SPESAYTAETT-----SPHPSHDGSSFKSPDTVCLSRGKLLVEKAIKDHDFIP------- 154
             :||    :       | | :|       ||||  | ||:        |||
Db     146 DDLTAYITGASLINQDFAHSSEEGLKKTEVITVCLIIGLLLIVFRSVVTPFIPIVVVGFS 205

Qy     155 ---SNSILSNALSWGVKILHID 173
               | |||      |: : ::|
Db     206 YLISQSIL------GILVYNVD 221
```

```
RESULT  10
TOP2_PEA
ID   TOP2_PEA       STANDARD;      PRT;  1462 AA.
AC   O24308;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DNA TOPOISOMERASE II (EC 5.99.1.3).
GN   TOP2 OR TOPII.
```

```
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Reddy M.K., Nair S., Tewari K.K.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS.
CC   -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC       OF DOUBLE-STRANDED DNA.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y14559; CAA74891.1; -.
DR   InterPro; IPR000947; CBFA_NFYB.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR003594; HATPase_c.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 2.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   PRINTS; PR00615; CCAATSUBUNTA.
DR   PRINTS; PR01158; TOPISMRASEII.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP-binding.
FT   NP_BIND     149    154       ATP (POTENTIAL).
FT   ACT_SITE    761    761       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   1462 AA;  164205 MW;  D9212C54AE0F8B2E CRC64;


  Query Match             7.2%;  Score 86.5;  DB 1;  Length 1462;
  Best Local Similarity   20.7%;  Pred. No. 42;
  Matches   53;  Conservative   30;  Mismatches   88;  Indels   85;  Gaps   12;

Qy      21 KNEKNRPSLKSLKTDNRPEKSKCKP------------LWGKVFYLDLPSVTI-------- 60
           |  :  :| :     |:  |:   :|             ||   | || |: |
Db    1044 KGKSTKPQVAGANDDDSEEQEDAEPETASQSVSVEGATWGD--YDDLLSLPIGTLTLESV 1101

Qy      61 ----SEKLQKD---------------IKDLGGRVEEFLSK----DISYLISNKKEAK--- 94
               || :|:                 :|||    :||  |    |: |   ::| |
Db    1102 QKLLDEKTEKEKEYEILSGTPTTSLWLKDL----DEFEKKLDELDLKYAEDDRKRASQGS 1157

Qy      95 -----FAQTLGRISPVPSPESAYTAETTSPHPSHDGSSFK--------SPDTVCLSRGKL 141
                ||    :  | |      | :  |  | :| || :          |  |   :|:
Db    1158 KKANGFASKPAKKPPQPRKN---TKKAKSVEPENDNSSMEIENAVEAAKPAEVAKPKGRA 1214

Qy     142 LVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYYIEQKKKELYLLKKSSTSVRDGGK 201
             :| |:     | : | |          |      | ||   ::   ::      :  ||
Db    1215 APKKNIQKE---PEDDIQS---------LQERLAAYNIESSGEKSQAMESEEVQQKAAGK 1262

Qy     202 RVGSGAQKTRTGRLKK 217
           :     |  : |  ||
Db    1263 K-----QNNKRGGAKK 1273
```

```
RESULT  11
TP6B_ARCFU
ID   TP6B_ARCFU     STANDARD;      PRT;   602 AA.
AC   O29605;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   TYPE II DNA TOPOISOMERASE VI SUBUNIT B (EC 5.99.1.3).
GN   TOP6B OR AF0652.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC   Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- FUNCTION: RELAXES BOTH POSITIVE AND NEGATIVE SUPERTURNS AND
CC       EXHIBITS A STRONG DECATENASE ACTIVITY. THE B SUBUNIT BINDS ATP (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC       OF DOUBLE-STRANDED DNA.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO SUBUNITS A AND TWO SUBUNITS B (BY
CC       SIMILARITY).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001060; AAB90588.1; -.
DR   TIGR; AF0652; -.
DR   InterPro; IPR000410; Bctrl_sensor.
DR   InterPro; IPR003594; HATPase_c.
DR   Pfam; PF02518; HATPase_c; 1.
DR   SMART; SM00387; HATPase_c; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP-binding; Complete proteome.
SQ   SEQUENCE   602 AA;  67399 MW;  FBA2111CF4280DC7 CRC64;


  Query Match             7.0%;  Score 85;  DB 1;  Length 602;
  Best Local Similarity   22.4%;  Pred. No. 17;
  Matches   47;  Conservative   36;  Mismatches   85;  Indels   42;  Gaps    9;

Qy      20 VKNEKNRPS-----LKSLKTDNRPEKSKCKPLWGKVFYLDLPSVTISEKLQKDIKDLGGR 74
           |:|    ||      |  | : | |  |: |      :     | :    :|   ::::| |
Db    396 VQNRGELPSAPAVILIHLASTNIPYTSESKESVAAI-----PEIIDETRLA--LQEVGRR 448

Qy      75 VEEFLSKDISYLISNKKEAKFAQTLGRISPVPSPESAYTAETTSPHPSHDGSSFKSPDTV 134
           ::|:| :        |||    : :|:: |: : :     |     |   :    |
Db    449 LKEYLERKSRQQKKKKE----EMIGKVLPLIAKK---VCEILEKEP------LEIDRIV 495

Qy     135 CLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYYIEQKKKELYLLKKSST 194
             | | ||: ::: |              |||:: |    :   : || : | :  |
```

```
Db    496 ARIMGYLHVERIVEERD--------------GVKVVTIRVSNF--TRSKKSIKLYEMCSG 539

Qy    195 SVRDGGKRVGSGAQKTRTGRLK-KPFVKVE 223
          :|   | :|     | |  |: ||  :||
Db    540 NVEADGAKVSGSGYSTVTWSLEVKPDEEVE 569


RESULT  12
SM30_RAT
ID   SM30_RAT       STANDARD;      PRT;   299 AA.
AC   Q03336; Q63496;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   SENESCENCE MARKER PROTEIN-30 (SMP-30) (REGUCALCIN) (RC).
GN   RGN OR SMP30.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=93041931; PubMed=1420310;
RA   Fujita T., Shirasawa T., Uchida K., Maruyama N.;
RT   "Isolation of cDNA clone encoding rat senescence marker protein-30
RT   (SMP30) and its tissue distribution.";
RL   Biochim. Biophys. Acta 1132:297-305(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=93351639; PubMed=8348951;
RA   Shimokawa N., Yamaguchi M.;
RT   "Molecular cloning and sequencing of the cDNA coding for a calcium-
RT   binding protein regucalcin from rat liver.";
RL   FEBS Lett. 327:251-255(1993).
RN   [3]
RP   SEQUENCE OF 117-299 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX   MEDLINE=96150990; PubMed=8569761;
RA   Shimokawa N., Matsuda Y., Yamaguchi M.;
RT   "Genomic cloning and chromosomal assignment of rat regucalcin gene.";
RL   Mol. Cell. Biochem. 151:157-163(1995).
RN   [4]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=96386712; PubMed=8794449;
RA   Fujita T., Shirasawa T., Uchida K., Maruyama N.;
RT   "Gene regulation of senescence marker protein-30 (SMP30): coordinated
RT   up-regulation with tissue maturation and gradual down-regulation
RT   with aging.";
RL   Mech. Ageing Dev. 87:219-229(1996).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
CC       IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC       CALCIUM SIGNALING IN THE AGED LIVER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- TISSUE SPECIFICITY: HEPATOCYTES AND RENAL PROXIMAL TUBULAR
CC       EPITHELIUM.
CC   -!- DEVELOPMENTAL STAGE: IN LIVER, THE FIRST PEAK OF EXPRESSION WAS
CC       FOUND IN 5-DAY-OLD NEONATES. EXPRESSION INCREASES FROM DAY 7 AND
CC       REACHES A PLATEU AT DAY 10. 3-6.5 MOTH-OLD ADULTS EXPRESS ABOUT A
CC       THIRD THE AMOUNT OF NEONATES LEVEL. IN KIDNEY, EXPRESSION
CC       INCREASES FROM DAY 21 AND REACHES A MAXIMAL LEVEL AT DAY 35,
CC       REMAINS HIGH UNTIL 3 MONTHS OF AGE.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69021; CAA48786.1; -.
DR   EMBL; D38467; BAA07490.1; -.
DR   EMBL; D31662; BAA06507.1; -.
DR   PIR; S27203; S27203.
KW   Calcium-binding.
FT   CONFLICT    148    148       N -> D (IN REF. 1).
SQ   SEQUENCE   299 AA;  33389 MW;  E3CF6D3FFCAE4E98 CRC64;


  Query Match             7.0%;  Score 84.5;  DB 1;  Length 299;
  Best Local Similarity   21.7%;  Pred. No. 7.6;
  Matches   46;  Conservative   37;  Mismatches   60;  Indels   69;  Gaps   15;

Qy     39 EKSKC--KPLWGK----VFYLDLPSVT------ISEKLQK----------DIKDLGGRVE 76
          | :|  |:| :    : ::|:|| |       || ::|:           ::  || |
Db     12 ENYRCGESPVWEEASKCLLFVDIPSKTVCRWDSISNRVQRVGVDAPVSSVALRQSGGYVA 71

Qy     77 EFLSK-------DISYLI-----SNKKEAKFAQTLGRISPVPSPESAYTAETTSPH--PS 122
             :|       | |  |       :||  :|    |:: |     :   || |:|
Db     72 TIGTKFCALNWEDQSVFILAMVDEDKKNNRFND--GKVDPAGRYFAGTMAEETAPAVLER 129

Qy    123 HDGSSFKSPDTVCLSRGKLLVEKAIKDHDFIPSNSI-LSNALSWGV--KIL-HIDDIRYY 178
          | || :            |  : ::| :     | : :|| | | :  ||  :|| : |
Db    130 HQGSLY-----------SLFPDHSVKKY----FNQVDISNGLDWSLDHKIFYYIDSLSYT 174

Qy    179 IE----------QKKKELYLLKKSSTSVRDG 199
          ::           :: :| ::|    : ||
Db    175 VDAFDYDLPTGQISNRRTVYKMEKDE-QIPDG 205


RESULT  13
RPGR_HUMAN
ID   RPGR_HUMAN     STANDARD;      PRT;   815 AA.
AC   Q92834; Q93039; O00737; O00702;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   X-LINKED RETINITIS PIGMENTOSA GTPASE REGULATOR.
GN   RPGR OR RP3 OR XLRP3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96241570; PubMed=8673101;
RA   Meindl A., Dry K., Herrmann K., Manson F., Ciccodicola A.,
RA   Edgar A., Carvalho M.R.S., Achatz H., Hellebrand H., Lennon A.,
RA   Migliaccio C., Porter K., Zrenner E., Bird A., Jay M., Lorenz B.,
RA   Wittwer B., D'Urso M., Meitinger T., Wright A.;
RT   "A gene (RPGR) with homology to the RCC1 guanine nucleotide exchange
RT   factor is mutated in X-linked retinitis pigmentosa (RP3).";
RL   Nat. Genet. 13:35-42(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RA   Berger W.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-509 FROM N.A., AND VARIANTS RP3 C-130; S-235 AND S-275.
RC   TISSUE=Retina;
RX   MEDLINE=96414315; PubMed=8817343;
RA   Roepman R., van Duijnhoven G., Rosenberg T., Pinckers A.J.L.G.,
RA   Bleeker-Wagemakers L.M., Bergen A.A.B., Post J., Beck A.,
RA   Reinhardt R., Ropers H.-H., Cremers F., Berger W.;
RT   "Positional cloning of the gene for X-linked retinitis pigmentosa 3:
RT   homology with the guanine-nucleotide-exchange factor RCC1.";
RL   Hum. Mol. Genet. 5:1035-1041(1996).
RN   [4]
RP   VARIANTS RP3.
RX   MEDLINE=98153625; PubMed=9399904;
RA   Buraczynska M., Wu W., Fujita R., Buraczynska K., Phelps E.,
```

```
RA   Andreasson S., Bennett J., Birch D.G., Fishman G.A., Hoffman D.R.,
RA   Inana G., Jacobson S.G., Musarella M.A., Sieving P.A., Swaroop A.;
RT   "Spectrum of mutations in the RPGR gene that are identified in 20% of
RT   families with X-linked retinitis pigmentosa.";
RL   Am. J. Hum. Genet. 61:1287-1292(1997).
RN   [5]
RP   VARIANT RP3 VAL-60.
RX   MEDLINE=99070804; PubMed=9855162;
RA   Fishman G.A., Grover S., Jacobson S.G., Alexander K.R., Derlacki D.J.,
RA   Wu W., Buraczynska M., Swaroop A.;
RT   "X-linked retinitis pigmentosa in two families with a missense
RT   mutation in the RPGR gene and putative change of glycine to valine at
RT   codon 60.";
RL   Ophthalmology 105:2286-2296(1998).
CC   -!- FUNCTION: COULD BE GUANINE-NUCLEOTIDE RELEASING FACTOR.
CC   -!- SUBCELLULAR LOCATION: POSSIBLY MEMBRANE-ANCHORED.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SMALL FORM
CC       LACKS PART OF REPEAT 7.
CC   -!- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, MUSCLE,
CC       KIDNEY, PANCREAS AND FETAL RETINAL PIGMENT EPITHELIUM.
CC   -!- DISEASE: DEFECTS IN RPGR ARE RESPONSIBLE FOR X-LINKED RETINITIS
CC       PIGMENTOSA-3 (XLRP-3 OR RP3), A FORM OF CHOROIDO-RETINAL
CC       DEGENERATION WHICH IS DISTINGUISHED FROM OTHER TYPES BY THE
CC       PRESENCE IN HETEROZYGOUS WOMEN OF A TAPETAL-LIKE RETINAL REFLEX (A
CC       BRILLIANT, SCINTILLATING, GOLDEN-HUED, PATCHY APPEARANCE MOST
CC       STRIKING AROUND THE MACULA) BUT NO VISUAL DEFECT.
CC   -!- SIMILARITY: CONTAINS 7 RCC1 REPEATS.
CC   -!- DATABASE: NAME=Mutations of the RPGR gene;
CC       NOTE=Retina International's Scientific Newsletter;
CC       WWW="http://www.retina-international.com/sci-news/rpgrmut.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U57629; AAC50481.1; -.
DR   EMBL; X97668; CAA66258.1; -.
DR   MIM; 312610; -.
DR   InterPro; IPR000408; RCC1.
DR   Pfam; PF00415; RCC1; 7.
DR   PROSITE; PS00625; RCC1_1; FALSE_NEG.
DR   PROSITE; PS00626; RCC1_2; 4.
DR   PROSITE; PS50012; RCC1_3; 6.
KW   Guanine-nucleotide releasing factor; Alternative splicing; Repeat;
KW   Vision; Retinitis pigmentosa; Disease mutation.
FT   REPEAT       11     52       RCC1 1.
FT   REPEAT       53    104       RCC1 2.
FT   REPEAT      106    157       RCC1 3.
FT   REPEAT      158    207       RCC1 4.
FT   REPEAT      208    260       RCC1 5.
FT   REPEAT      262    312       RCC1 6.
FT   REPEAT      314    366       RCC1 7.
FT   VARSPLIC    354    415       MISSING (IN SHORT ISOFORM).
FT   VARIANT      60     60       G -> V (IN RP3).
FT                                /FTId=VAR_008501.
FT   VARIANT      60     60       G -> N (IN RP3).
FT                                /FTId=VAR_008502.
FT   VARIANT      75     75       I -> V (IN RP3).
FT                                /FTId=VAR_008503.
FT   VARIANT      98     98       H -> Q (IN RP3).
FT                                /FTId=VAR_008504.
FT   VARIANT     130    130       F -> C (IN RP3).
FT                                /FTId=VAR_006850.
FT   VARIANT     215    215       G -> V (IN RP3).
FT                                /FTId=VAR_008505.
FT   VARIANT     235    235       P -> S (IN RP3).
FT                                /FTId=VAR_006851.
FT   VARIANT     250    250       C -> R (IN RP3).
FT                                /FTId=VAR_008506.
FT   VARIANT     262    262       A -> G (IN RP3).
FT                                /FTId=VAR_008507.
FT   VARIANT     275    275       G -> S (IN RP3).
FT                                /FTId=VAR_006852.
FT   VARIANT     425    425       R -> K (IN RP3).
FT                                /FTId=VAR_008508.
FT   VARIANT     431    431       I -> V (IN RP3).
FT                                /FTId=VAR_008509.
FT   VARIANT     436    436       G -> D (IN RP3).
FT                                /FTId=VAR_008510.
FT   VARIANT     566    566       G -> E (IN RP3).
FT                                /FTId=VAR_008511.
FT   CONFLICT      1      3       MRE -> MAKLRRSTTTAL (IN REF. 3).
SQ   SEQUENCE   815 AA;  90244 MW;  70D84EAD988348D1 CRC64;


  Query Match             7.0%;  Score 84;  DB 1;  Length 815;
  Best Local Similarity   20.0%;  Pred. No. 31;
  Matches   48;  Conservative   40;  Mismatches   78;  Indels   74;  Gaps    8;

Qy      20 VKNEKNRPSLKSLKTD----------------NRPEKSKCKPLWGKVFYLDLPSVTIS-- 61
           |: :| : ::  |  |                    |   ||    :  ::  |: ||
Db     520 VQKQKKQQTIGELTQDTALTENDDSDEYEEMSEMKEGKACKQHVSQGIFMTQPATTIEAF 579

Qy      62 --------------------------EKLQKDIKDLGGRVE--EFLSKDIS-----YLI 87
                                       |  ::::|  ||| |  | || |::     :
Db     580 SDEEVEIPEEKEGAEDSKGNGIEEQEVEANEENVKVHGGRKEKTEILSDDLTDKAEDHEF 639

Qy      88 SNKKEAKF---------------AQTLGRISPVPSPESAYTAETTSPHPSHDGSSFKSPD 132
           |  :| |                 :|:|    ||   : | ::|     ::| |   | :
Db     640 SKTEELKLEDVDEEINAENVESKKKTVGDDESVP---TGYHSKTEGAERTNDDS---SAE 693

Qy     133 TVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYYIEQKK-KELYLLKK 191
           |:       | |:|| :::  |   :| :| |  ::||   :     : || |:::| |:
Db     694 TIEKKEKANLEERAICEYNENPKGYMLDDADSSSLEILENSETTPSKDMKKTKKIFLFKR 753


RESULT  14
ADVL_HUMAN
ID   ADVL_HUMAN     STANDARD;      PRT;   819 AA.
AC   O75366;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   ADVILLIN (P92).
GN   AVIL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RX   MEDLINE=98330437; PubMed=9664034;
RA   Marks P.W., Arai M., Bandura J.L., Kwiatkowski D.J.;
RT   "Advillin (p92): a new member of the gelsolin/villin family of actin
RT   regulatory proteins.";
RL   J. Cell Sci. 111:2129-2136(1998).
CC   -!- FUNCTION: CA(2+)-REGULATED ACTIN-BINDING PROTEIN.
CC   -!- TISSUE SPECIFICITY: MOST HIGHLY EXPRESSED IN THE SMALL INTESTINE
CC       AND COLONIC LINING. WEAKER EXPRESSION ALSO DETECTED IN THE THYMUS,
CC       PROSTATE, TESTES AND UTERUS.
CC   -!- SIMILARITY: BELONGS TO THE VILLIN/GELSOLIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 6 GELSOLIN-LIKE REPEATS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; AF041449; AAC25051.1; -.
DR   HSSP; P02640; 1VII.
DR   InterPro; IPR001974; Gelsolin.
DR   InterPro; IPR003128; VHP.
DR   Pfam; PF00626; Gelsolin; 6.
DR   Pfam; PF02209; VHP; 1.
DR   PRINTS; PR00597; GELSOLIN.
DR   SMART; SM00262; GEL; 6.
DR   SMART; SM00153; VHP; 1.
KW   Cytoskeleton; Calcium; Actin-binding; Capping protein; Repeat.
FT   DOMAIN        1    731       CORE (BY SIMILARITY).
FT   DOMAIN      731    819       HEADPIECE (BY SIMILARITY).
FT   REPEAT       24     73       GELSOLIN-LIKE 1.
FT   REPEAT      145    185       GELSOLIN-LIKE 2.
FT   REPEAT      262    306       GELSOLIN-LIKE 3.
FT   REPEAT      403    454       GELSOLIN-LIKE 4.
FT   REPEAT      525    565       GELSOLIN-LIKE 5.
FT   REPEAT      628    669       GELSOLIN-LIKE 6.
FT   DOMAIN      784    819       VHP.
FT   SITE        109    116       POLYPHOSPHOINOSITIDE BINDING (BY
FT                                SIMILARITY).
FT   SITE        135    143       POLYPHOSPHOINOSITIDE BINDING (BY
FT                                SIMILARITY).
SQ   SEQUENCE   819 AA;  92086 MW;  8EE1C1ED60110CB7 CRC64;


  Query Match             7.0%;  Score 84;  DB 1;  Length 819;
  Best Local Similarity   22.7%;  Pred. No. 32;
  Matches   61;  Conservative   38;  Mismatches   76;  Indels   94;  Gaps   16;

Qy      11 KGHFQGGIQVKNEKNRPSLKSLKTDNRPEK-------------SKCKPLW-----GKVFY 52
           :|:|: ||  |       :| ::|:    |              :: :  |     | ||
Db     105 RGYFKQGIIYKQGGVASGMKHVETNTYDVKRLLHVKGKRNIRATEVEMSWDSFNRGDVFL 164

Qy      53 LDLPSVTI---------SEKLQ-----KDIKD--LGGRVEEFLSKDISYLISNKKEAK-- 94
           |||  | |          |:|:     |||:|   ||| |      |  :  :|: |
Db     165 LDLGKVIIQWNGPESNSGERLKAMLLAKDIRDRERGGRAE------IGVIEGDKEAASPE 218

Qy      95 ----FAQTLGRISPVPSPESAYTAETTSPHPSHDGSSFKSPDTVCL-----SRGKLLVEK 145
                |||| | :          : | |     |    |   |: |     | |:| | :
Db     219 LMKVLQDTLGRRSII---------KPTVPDEIIDQ---KQKSTIMLYHISDSAGQLAVTE 266

Qy     146 AIKDHDFIPSNSILSNALSWGVKILHIDDIRYYIEQKKKELYLLKKSSTSVRDGGKRVGS 205
                : :  :: :       :|: ||  | ::|   ::|: |              |
Db     267 -------VATRPLVQD-------LLNHDDC-YILDQSGTKIYVWK-------------GK 298

Qy     206 GAQKT-RTGRLKKP--FVKVEDMSQSPAV 231
           || |  :   : |   |:|::    |  |
Db     299 GATKAEKQAAMSKALGFIKMKSYPSSTNV 327


RESULT  15
CERU_RAT
ID   CERU_RAT       STANDARD;      PRT;  1059 AA.
AC   P13635; Q64719;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE).
GN   CP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung, and Liver;
RX   MEDLINE=90237081; PubMed=2332446;
RA   Fleming R.E., Gitlin J.D.;
RT   "Primary structure of rat ceruloplasmin and analysis of
RT   tissue-specific gene expression during development.";
RL   J. Biol. Chem. 265:7701-7707(1990).
RN   [2]
RP   SEQUENCE OF 257-294; 571-612 AND 823-892 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=87137545; PubMed=3818625;
RA   Aldred A.R., Grimes A., Schreiber G., Mercer J.F.B.;
RT   "Rat ceruloplasmin. Molecular cloning and gene expression in liver,
RT   choroid plexus, yolk sac, placenta, and testis.";
RL   J. Biol. Chem. 262:2875-2878(1987).
CC   -!- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER
CC       MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS
CC       ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT
CC       AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
CC   -!- FUNCTION: MAY ALSO PLAY A ROLE IN FETAL LUNG DEVELOPMENT OR
CC       PULMONARY ANTIOXIDANT DEFENSE.
CC   -!- CATALYTIC ACTIVITY: 4 FE(2+) + 4 H(+) + O(2) = 4 FE(3+) + 2 H(2)O.
CC   -!- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO
CC       THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS
CC       KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED
CC       BINUCLEAR.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED INTO THE
CC       PLASMA. ALSO CHOROID PLEXUS, YOLK SAC, PLACENTA, AND TESTIS; NOT
CC       IN STOMACH AND SMALL INTESTINE. FETAL LUNG AND LIVER.
CC   -!- INDUCTION: BY INFLAMMATION.
CC   -!- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC       2 PLASTOCYANIN-LIKE REPEATS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L33869; AAA40917.1; -.
DR   EMBL; M80529; AAB65820.1; -.
DR   EMBL; J02670; AAA40914.1; ALT_SEQ.
DR   EMBL; M14102; AAA40915.1; -.
DR   PIR; A35210; A35210.
DR   PIR; A29564; A29564.
DR   HSSP; P00450; 1KCW.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR002355; MultiCu_oxidse2.
DR   Pfam; PF00394; Cu-oxidase; 3.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR   PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW   Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;
KW   Signal.
FT   SIGNAL        1     19       PROBABLE.
FT   CHAIN        20   1059       CERULOPLASMIN.
FT   DOMAIN       20    356       F5/8 TYPE A 1.
FT   DOMAIN       20    199       PLASTOCYANIN-LIKE 1.
FT   DOMAIN      208    354       PLASTOCYANIN-LIKE 2.
FT   DOMAIN      369    712       F5/8 TYPE A 2.
FT   DOMAIN      369    554       PLASTOCYANIN-LIKE 3.
FT   DOMAIN      564    710       PLASTOCYANIN-LIKE 4.
FT   DOMAIN      724   1055       F5/8 TYPE A 3.
FT   DOMAIN      724    894       PLASTOCYANIN-LIKE 5.
FT   DOMAIN      902   1051       PLASTOCYANIN-LIKE 6.
FT   DISULFID    173    199       BY SIMILARITY.
FT   DISULFID    275    356       BY SIMILARITY.
FT   DISULFID    528    554       BY SIMILARITY.
FT   DISULFID    631    712       BY SIMILARITY.
FT   DISULFID    868    894       BY SIMILARITY.
FT   METAL       120    120       COPPER (TYPE 2) (BY SIMILARITY).
FT   METAL       122    122       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL       179    179       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL       181    181       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL       988    988       COPPER (TYPE 1) (BY SIMILARITY).
FT   METAL       991    991       COPPER (TYPE 2) (BY SIMILARITY).
FT   METAL       993    993       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL      1033   1033       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL      1034   1034       COPPER (TYPE 1) (BY SIMILARITY).
```

```
FT   METAL      1035   1035       COPPER (TYPE 3) (BY SIMILARITY).
FT   METAL      1039   1039       COPPER (TYPE 1) (BY SIMILARITY).
FT   METAL      1044   1044       COPPER (TYPE 1) (BY SIMILARITY).
FT   CARBOHYD    138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    582    582       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    756    756       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    920    920       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    271    271       G -> A (IN REF. 2).
FT   CONFLICT    604    605       ED -> DN (IN REF. 2).
FT   CONFLICT    823    823       T -> S (IN REF. 2).
FT   CONFLICT    833    833       V -> L (IN REF. 2).
FT   CONFLICT    868    868       C -> V (IN REF. 2).
FT   CONFLICT    891    891       L -> R (IN REF. 2).
SQ   SEQUENCE   1059 AA;  120840 MW;  12BA3B990A0B95E3 CRC64;


  Query Match             7.0%;  Score 84;  DB 1;  Length 1059;
  Best Local Similarity   20.8%;  Pred. No. 44;
  Matches   44;  Conservative   29;  Mismatches   83;  Indels   56;  Gaps    9;

Qy      10 SKGHFQGGIQVKNEKNRPSLKSLKTDNRPEKSKCKPLW----GKVFYLDLPSVTISEKLQ 65
           |  : : |      | : :| |  ||    |:  || |    | |         : :|:
Db      52 SNFYLRNGPDRIGRKYKKALYSEYTDGTFTKTIDKPAWLGFLGPVI-----KAEVGDKVS 106

Qy      66 KDIKDLGGRVEEFLSKDISYLISNKKEAKFAQTLGRISP------------VPSPESAY 112
             :|:   |   | :  ::|  :|:         | | |             |  :  |
Db     107 VHVKNFASRPYTFHAHGVTYTKANE---------GAIYPDNTTDFQRADDKLFPGQQYLY 157

Qy     113 TAETTSPHP------------SH------DGSSFKSPDTVCLSRGKLLVEKAIK-DHDF 152
             | |             ||        |    |  :|  :| |  ||    | :|
Db     158 VLRANEPSPGEGDSNCVTRIYHSHVDAPKDIASGLIGPLILC-KKGSLHKEKEENIDQEF 216

Qy     153 IPSNSILSNALSWGVKILHIDDIRYYIEQKKK 184
           :   |::   ||| ::    |:|: :  : :|
Db     217 VLMFSVVDENLSWYLE----DNIKTFCSEPEK 244


Search completed: December 27, 2001, 16:56:15
Job time: 230 sec
```

THIS PAGE BLANK (USPTO)